Automatica 50 (2014) 1160–1166

Contents lists available at ScienceDirect

Automatica

journal homepage: www.elsevier.com/locate/automatica

Brief paper

Identification of smooth nonlinear dynamical systems with non-smooth steady-state features☆

Luis A. Aguirre [1]

Departamento de Engenharia Eletrônica, Universidade Federal de Minas Gerais, Av. Antonio Carlos 6627, 31.270-901 Belo Horizonte, MG, Brazil

ARTICLE INFO

Article history:
Received 8 January 2013
Received in revised form
28 October 2013
Accepted 23 December 2013
Available online 25 February 2014

Keywords:
Nonlinear system identification
Nonlinear models
Gray-box identification
Bifurcation analysis
Dead zones

ABSTRACT

This work discusses the identification of single-block smooth nonlinear discrete-time polynomial models with non-smooth steady-state features. Based on bifurcation theory, conditions are developed and used to determine some general aspects of the model structure and also to determine some constraints on the parameters required to guarantee the aforementioned features. The procedure uses only smooth functions of the regressors, a single possibly smooth input and some prior knowledge about the steady-state behavior. The non-smooth static function is here obtained by interchanging the stability of two sets of equilibria at the break-point, which corresponds to guaranteeing a transcritical bifurcation. This work discusses how to determine the domain over which the results are valid. The procedure is illustrated with simulated and experimental data.

1. Introduction

Nonlinear system identification is a challenging problem, especially when the data are produced by a system with non-smooth static nonlinearities. This problem has been dealt with in different ways, as for instance using smooth static functions to approximate the non-smooth static function (Merzouki, Davila, Fridman, & Cadiou, 2007). A workable approach uses block-oriented models, such as the Hammerstein (Bai, 2002; Giri, Rochdi, & Chaoui, 2009) or the sandwich (Tan, Dong, & Li, 2009) representations. In such cases non-smooth and even discontinuous static functions are approximated by non-smooth algebraic functions in order to accurately represent discontinuities at the break-points. Specific control schemes assume that an algebraic function (usually non-smooth) is used to approximate the static characteristic (Zhou, Wen, & Zhang, 2006), whereas the dynamics are modeled by a (smooth) differential or difference equation. Some nonlinear control schemes assume that the system is modeled by a single-block model, typically a NARX (nonlinear autoregressive model with exogenous inputs) (Napoli & Piroddi, 2010; Pröll & Karim, 1994). In such cases, the inclusion of non-smooth functions to model the static characteristic would also affect the dynamics, which are usually smooth. The identification of systems with hysteresis has been discussed in Leva and Piroddi (2002) where additional non-smooth inputs computed from the data have been employed and actually play the role of non-smooth regressors. The apparent incompatibility of having single-block models with static functions that may have break-points seems to be confirmed by the lack of methods to identify such models with smooth dynamics yet with non-smooth steady-state features.

This work presents a gray-box procedure for the identification of single-block NARX polynomial models with smooth dynamics, which can accurately reproduce a class of steady-state functions with one break-point. This is achieved by designing two sets of equilibria, one for each part of the steady-state function (SSF) considered, and switching their stability at the desired break-point using results derived from bifurcation theory. Some aspects of the SSF are considered to be known a priori. The importance of using prior knowledge in modeling and control has been argued in Bars et al. (2008). The conditions under which such bifurcation takes place *for NARX polynomial models* and the domain of validity of the results are given using simple lemmas. The ease with which the lemmas are applied and understood will enable tackling the difficult problem at hand. The key contribution and novelty of this work lays in the adaptation of bifurcation theory to NARX polynomial models in order to achieve a non-smooth steady-state behavior with a model composed solely of smooth functions. This paper deals with the transcritical bifurcation which is useful to

☆ This work has been partially supported by CNPq and FAPEMIG. The material in this paper was not presented at any conference. This paper was recommended for publication in revised form by Associate Editor Martin Enqvist under the direction of Editor Torsten Söderström.

E-mail address: aguirre@cpdee.ufmg.br.

[1] Tel.: +55 31 3409 4866; fax: +55 31 3409 4850.

http://dx.doi.org/10.1016/j.automatica.2014.02.012

model the steady-state behavior of some control valves. Other bifurcations have been addressed elsewhere (Aguirre & Furtado, 2007; Rodrigues & Aguirre, 2012).

The work is organized as follows. Section 2 provides background material. The problem is stated in Section 3 and solved in Section 4 for a class of steady-state functions and in Section 5 for a slightly more general class. Section 6 provides an example with experimental data. The main conclusions of the paper are given in Section 7.

2. Background

2.1. Nonlinear polynomial models

A general NARX model (nonlinear autoregressive model with exogenous inputs) can be written as (Billings, Chen, & Korenberg, 1989)

$$y(k) = F[y(k-1), \dots, y(k-n_y),\ u(k-d) \dots u(k-n_u)] + e(k), \tag{1}$$

where $u(k)$ and $y(k)$ are respectively the input and output signals and $e(k)$ accounts for uncertainties and possible noise. $n_y,\ n_u,\ d \in \mathbb{N}$ are the maximum lags for output, input and the time delay, respectively. In this work $F[\cdot]$ is assumed to be a polynomial with nonlinearity degree ℓ. In prediction error minimization (PEM) estimation problems, a moving average (MA) part is usually included in the model to reduce bias. In this paper constrained parameter estimation will be pursued and therefore the MA part will not be used. The NARX part of model (1) can be expanded as the summation of terms with degrees of nonlinearity in the range $[1\ \ \ell]$. Each $(p+m)$th-degree term can contain a pth-degree factor in y and an mth-degree factor in u and is multiplied by a constant parameter $c_{p,m}(\tau_1, \dots, \tau_{p+m})$ as follows

$$y(k) = \sum_{m=0}^{\ell} \sum_{p=0}^{\ell-m} \sum_{\tau_1,\tau_m}^{n_y,n_u} c_{p,m}(\tau_1, \dots, \tau_{p+m}) \prod_{j=1}^{p} y(k-\tau_j) \times \prod_{i=1}^{m} u(k-\tau_{p+i}) + e(k), \tag{2}$$

where the third summation is

$$\sum_{\tau_1,\tau_m}^{n_y,n_u} \equiv \sum_{\tau_1=1}^{n_y} \cdots \sum_{\tau_p=1}^{n_y} \sum_{\tau_{p+1}=d}^{n_u} \cdots \sum_{\tau_{p+m}=d}^{n_u}, \tag{3}$$

and the upper limit is n_y if the summation refers to factors in y or n_u for factors in u. The model structure can be chosen using orthogonal techniques (Billings et al., 1989).

The choice of which terms to use in order to compose the final model is known as the structure selection problem. Effective solutions to this problem include the error reduction ratio (ERR) (Billings et al., 1989) and the simulation reduction ratio (SRR) criterion (Piroddi, 2008). See Hong et al. (2008) for a review of structure selection methods.

Steady-state analysis is accomplished by taking $\bar{y} = y(k-\tau_j),\ \forall \tau_j = 1, \dots, n_y,\ \bar{u} = u(k-\tau_i),\ \forall \tau_i = d, \dots, n_u$ and in that case Eq. (2) can be rewritten as

$$\bar{y} = \sum_{m=0}^{\ell} \sum_{p=0}^{\ell-m} \left(\sum_{\tau_1,\tau_m}^{n_y,n_u} c_{p,m}(\tau_1, \dots, \tau_{p+m}) \right) \bar{y}^p \bar{u}^m = \sum_{m=0}^{\ell} \sum_{p=0}^{\ell-m} \Sigma_{y^p u^m} \bar{y}^p \bar{u}^m. \tag{4}$$

The solution of (4) will yield the fixed points, or equilibria, of model (2) for a given $\bar{u}$.

The constant within the large parenthesis in Eq. (4), denoted $\Sigma_{y^p u^m}$, is the *cluster coefficient* of a set of model terms in (2), called *term cluster*, indicated by $\Omega_{y^p u^m}$. Terms of the form $y^p(k-\tau_j)u^m(k-\tau_i) \in \Omega_{y^p u^m}$ for $m+p \le \ell$, where τ_i and τ_j are any time lags (Aguirre & Billings, 1995).

For instance, for the model $y(k) = \theta_1 y(k-1)y(k-2) + \theta_2 y(k-1)u(k-2) + \theta_3 y(k-3)u(k-3)$ we have $n_y = n_u = 3,\ d = 2,\ \ell = 2$ and $\theta_1 = c_{2,0}(1, 2)$, $\theta_2 = c_{1,1}(1, 2)$ and $\theta_3 = c_{1,1}(3, 3)$. This model has two term clusters, namely: Ω_{y^2} with coefficient $\Sigma_{y^2} = \theta_1$ and Ω_{uy} with coefficient $\Sigma_{uy} = \theta_2 + \theta_3$. Hence, in steady-state (see Eq. (4)) $\bar{y} = \Sigma_{y^2}\bar{y}^2 + \Sigma_{uy}\bar{y}\bar{u}$.

2.2. Constrained parameter estimation

Model (2) can be written as $y(k) = \boldsymbol{\psi}^{\mathrm{T}}(k-1)\hat{\boldsymbol{\theta}} + \xi(k)$, where $\boldsymbol{\psi}(k-1)$ is the regressors vector which contains linear and nonlinear combinations of output and input terms up to and including time $k-1$ and $\xi(k)$ is the residual at instant k. Consider the set of constraints on the parameter vector written as $\mathbf{c} = S\boldsymbol{\theta}$, where $\mathbf{c}$ is a given constant vector, and S is a known constant matrix. The solution to the problem

$$\hat{\boldsymbol{\theta}}_{\mathrm{CLS}} = \underset{\boldsymbol{\theta}\,:\,\mathrm{c} = S\boldsymbol{\theta}}{\arg\min} \left[\boldsymbol{\xi}^{\mathrm{T}}\boldsymbol{\xi}\right], \tag{5}$$

is given by (Draper & Smith, 1998)

$$\hat{\boldsymbol{\theta}}_{\mathrm{CLS}} = \hat{\boldsymbol{\theta}}_{\mathrm{LS}} - (\Psi^{\mathrm{T}}\Psi)^{-1}S^{\mathrm{T}}[S(\Psi^{\mathrm{T}}\Psi)^{-1}S^{\mathrm{T}}]^{-1}(S\hat{\boldsymbol{\theta}}_{\mathrm{LS}} - \mathbf{c}), \tag{6}$$

where Ψ is the regressors matrix and $\hat{\boldsymbol{\theta}}_{\mathrm{LS}} = (\Psi^{\mathrm{T}}\Psi)^{-1}\Psi^{\mathrm{T}}\mathbf{y}$ is the standard least-squares solution. The estimator in (6) will be biased in general. However, the aim in this work is to be able to impose a transcritical bifurcation and this will be achieved at the expense of some bias.

2.3. The transcritical bifurcation

The normal form of the supercritical transcritical bifurcation is given by

$$y(k) = F_{\mathrm{t}}[y(k-1), \mu] = y(k-1) + \mu\, y(k-1) - y(k-1)^2, \tag{7}$$

where μ is the bifurcation parameter. The steady-state behavior of (7) is reached by taking $\bar{y} = y(k) = y(k-1)$, thus yielding

$$0 = \mu\,\bar{y} - \bar{y}^2 = \bar{y}(\mu - \bar{y}). \tag{8}$$

The fixed points or equilibria of Eq. (7) are given by the two lines $\bar{y} = \mu$ and $\bar{y} = 0$, that intersect at the origin of the $\bar{y} - \mu$ plane, as shown in Fig. 1(a). The conditions for stability are given by

$$-1 < \left.\frac{\partial F_{\mathrm{t}}}{\partial y(k-1)}\right|_{y(k-1)=\bar{y}} < 1$$
$$-1 < 1 + \mu - 2y(k-1)|_{y(k-1)=\bar{y}} < 1. \tag{9}$$

For the trivial fixed point $\bar{y} = 0$, conditions (9) becomes $-2 < \mu < 0$, and for the nontrivial fixed point $\bar{y} = \mu$, stability holds for $0 < \mu < 2$, as shown in Fig. 1(b).

3. Statement of the problem

The steady-state behavior of the NARX model (1) is obtained by taking $\bar{y} = y(k-\tau_j), \forall \tau_j = 1, \dots, n_y,\ \bar{u} = u(k-\tau_i), \forall \tau_i = d, \dots, n_u$ to yield

$$\bar{y} = \bar{F}[\bar{y}, \bar{u}]. \tag{10}$$

For a given constant value of the input, depending on the model function F, the output could have one or more possible values